

Applicant : Vincent P. Stanton, Jr.
Serial No. : 09/967,013
Filed : October 25, 2000
Page : 2

Attorney's Docket No.: 11926-022001 / 0015.UTL1

interrupted pallindrome recognition site like Mwo I (GCNNNNN/NNGC) (SEQ ID NO:1) can be positioned such that the first GC is in the primer while the second GC includes the polymorphic nucleotide. Only the allele corresponding to GC at the second site will be cleaved. Use of such restriction endonucleases simplifies the sequence requirements at and about the polymorphic site (in this example all that is required is that one allele at the polymorphic site include the dinucleotide GC), thereby increasing the number of polymorphic sites that can be analyzed in this way.--

Replace the paragraph beginning at page 26, line 6, with the following rewritten paragraph:

--**Table 2** Twenty polymorphic sites in the ApoE gene. The ApoE genomic sequence is taken from GenBank accession AB012576. The gene is composed of four exons and three introns. The transcription start site (beginning of first exon) is at nucleotide (nt) 18,371 of GenBank accession AB012576, while the end of the transcribed region (end of the 3' untranslated region, less polyA tract) is at nt 21958. The twenty polymorphic sites are depicted as shaded nucleotides in the Table, and are as follows (nucleotide position and possible nucleotides): 16541 (T/G); 16747 (T/G); 16965 (T/C); 17030 (G/C); 17098 (A/G); 17387 (T/C); 17785 (G/A); 17874 (T/A); 17937 (C/T); 18145 (G/T); 18476 (G/C); 19311 (A/G); 20334 (A/G); 21250 (C/T); 21349 (T/C); 21388 (T/C); 23524 (A/G); 23707 (A/C); 23759 (C/T); 23805 (G/C); and 37237 (G/A) (SEQ ID NO:5). The bold sequence listing indicates the transcribed sequence of the ApoE gene; the grey shaded region indicates the ApoE gene enhancer element; the underlined sequence depicts the coding region of the ApoE gene. Where polymorphisms result in a change of the amino acid sequence, the amino acid alteration is indicated, for example at nucleotide position 20334 the A/T polymorphism results in a alanine/threonine respectively at amino acid position 18 of the ApoE gene product. As described in the Detailed Description below, the polymorphisms at positions GenBank nucleotide number 17874, 17937, 18145, 18476, 21250, and 21388 have been previously described.--

Replace the paragraph beginning at page 29, line 18, with the following rewritten paragraph:

--**Figure 8** Shown in this figure is the incorporation of a single restriction enzyme recognition site in the amplicon for subsequent digestion and mass spectrometric analysis of the prepared fragments. Shown in this figure is incorporation of BcgI, an restriction enzyme that is capable of making two double strand cuts, one on the 5' side and one on the 3' side of their recognition site. The recognition site for BcgI is 12/10(N)CGA(N)₆TGC(N)12/10 (SEQ ID NO:2), which after digestion results in fragments sufficient for mass spectrometric analysis and identification of the polymorphic base with the fragment.--

Replace the paragraph beginning at page 29, line 26, with the following rewritten paragraph:

--**Figure 9** Shown in this figure is an example of the utility in the present invention of including a restriction enzyme recognition site for which the restriction enzyme creates a nick in the DNA amplicon instead of causing a double strand break. As shown in this figure, a primer R is designed to incorporate a N.BstNB I recognition site (GAGTCNNNN[^]NN) (SEQ ID NO:3) in addition to a FokI restriction site. As in previous figures, the primer forms a hair-pin loop structure when hybridized to the target DNA region, however, the PCR amplicon has the incorporated restriction site sequences. Digestion with FokI and N.BstNB I results in a 10 mer fragment that contains the polymorphic base (T in *italic*). Such a fragment is sufficient for analysis using a mass spectrometer.--

Replace the paragraph beginning at page 38, line 32, with the following rewritten paragraph:

--It may also only be necessary to incorporate one restriction enzyme site into the amplicon via the primer. This can be done if the enzyme utilized is capable of making two double strand cuts, one on the 5' side and one on the 3' side of the recognition site. An example of such an enzyme is Bcg I which has a recognition site of 12/10(N)CGA(N)₆TGC(N)12/10 (SEQ ID NO:2) (Figure 8). The arrows designate the sites of cleavage on both strands. This particular enzyme would generate fragments greater than the current optimal length for mass spec analysis, so similar enzymes that are capable of cleaving in a similar fashion but which would generate smaller fragments are more desirable. Also, as mass spectrometry techniques and

Applicant : Vincent P. Stanton, Jr.
Serial No. : 09/967,013
Filed : October 25, 2000
Page : 4

Attorney's Docket No.: 11926-022001 / 0015.UTL1

instrumentation for DNA analysis progress , it may be possible to reliably analyze DNA fragments of this length or greater obtaining the sensitivity and the resolution necessary to see single base differences in fragments of this length.--

Replace the paragraph beginning at page 39, line 6, with the following rewritten paragraph:

--Restriction enzymes can also be used which only nick the DNA instead of causing a double strand break. One such enzyme is N.BstNB I whose recognition site is GAGTCNNNN^NN (SEQ ID NO:3). The fragments generated by this scheme are outlined in Figure 9. This strategy would generate only one small fragment (10-mer in this case) instead of two which may make analysis less complicated, especially in an automated mode.--

Replace the paragraph beginning at page 64, line 1, with the following rewritten paragraph:

--Some of the alternative structures that would be useful include: (i) recognition sites for various DNA modifying enzymes such as restriction endonucleases, (ii) a cruciform DNA structure, that could be very stable, or could be recognized by enzymes such as bacteriophage resolvases (e.g., T4E7, T7EI), or (iii) recognition sites for DNA binding proteins (preferably from thermophilic organisms) such as zinc finger proteins, catalytically inactive endonucleases, or transcription factors. The purpose of inducing such structures in an allele specific manner would be to effect allele specific binding to, or modification of, DNA. For example, consider a duplex formed only (or preferentially) by a strand from one allele that contains the recognition sequence for a thermostable restriction enzyme such as Taq I. Allele specific strand cleavage could be achieved by inclusion of (thermostable) Taq I during the PCR, resulting in complete inactivation of each cleaved template molecule and thereby leading to allele selective amplification. What are the limits of such an approach? One requirement is that there are no Taq I sites elsewhere in the PCR amplicon; another is that one of the two alleles must form a Taq I recognition sequence. The former limitation - which would limit the length of amplicons that could be allele specifically modified if only frequently occurring restriction sites were used - as well as the latter limitation - which requires that polymorphisms occur in restriction endonuclease recognition sites - can be addressed in part by designing a 5' primer extension, along with an internal primer loop, so that the recognition sequence for a rare cutting restriction

Applicant : Vincent P. Stanton, Jr.
Serial No. : 09/967,013
Filed : October 25, 2000
Page : 5

Attorney's Docket No.: 11926-022001 / 0015.UTL1

endonuclease that (i) is an interrupted pallindrome, or (ii) cleaves at some distance from its recognition sequence is formed by the internal loop, while (i) the other end of the interrupted pallindrome, or (ii) the cleavage site for the restriction enzyme, occurs at the polymorphic nucleotide, and is therefore sensitive to whether there is a duplex or a (partially or completely) single stranded region at the polymorphic site. This scheme is illustrated in Figure 20. Preferred enzymes for PCR implementation of these schemes would include enzymes from thermophiles, such as Bsl I (CCNNNNN/NNGG) (SEQ ID NO:4) and Mwo I (GCNNNNN/NNGC) (SEQ ID NO:1).--